

系统综述方法与NIH实验可重复性倡议

Kristina A. Thayer¹, Mary S. Wolfe¹, Andrew A. Rooney¹, Abee L. Boyles¹, John R. Bucher¹, Linda S. Birnbaum²

¹Division of the National Toxicology Program, and ²Office of the Director, National Institute of Environmental Health Sciences, National Institutes of Health, Department of Health and Human Services, Research Triangle Park, North Carolina, USA
E-mail: thayer@niehs.nih.gov

在2005年发表于*PLoS Medicine*的一篇里程碑式的文章中, Ioannidis断定目前已发表的大部分研究结果有误 (Ioannidis 2005)。与此相映衬的是, 依据既往发表的临床前期研究结果而进行的药物开发和诸多临床试验进展不顺利或失败, 原因在于那些研究结果不能被进一步的研究证实。美国环境卫生科学研究所 (NIH) 最近针对研究结果无法重复验证的问题提出了一系列改进倡议 (Collins and Tabak 2014)。本文将概述目前为解决研究可重复性问题而进行的努



Kristina A. Thayer



Mary S. Wolfe



Andrew A. Rooney



Abee L. Boyles



John R. Bucher



Linda S. Birnbaum

力, 同时倡议使用系统综述方法来改善研究的可重复性。发展该方法的初衷在于评估用于公共卫生决策的研究证据质量的可信度。

越来越多的生物医学领域的报告和评论开始关注已发表科学发现难以被重复验证的问题。虽然对研究方法关键点不够完善的报告是导致该问题的原因之一, 其他因素如研究的执行也同样重要, 甚至更为重要 (Begley and Ellis 2012; Ioannidis 2005; Landis et al. 2012; Tsilidis et al. 2013)。这种情况已经促使私营和公共部门采取措施, 其中包括*Plos One* (<http://www.plosone.org/>), Science Exchange (<https://www.scienceexchange.com/>), figshare (<http://figshare.com/>)和Mendeley (<http://www.mendeley.com/>) (Nice 2013; Wadman 2013) 之间开展的非政府间的可重复性合作计划。该计划尝试重复2010至2012年间肿瘤生物学领域发表的50项最有影响力研究的关键发现。另一项重要的公共措施是NIH发起的“增强研究发现的重复性和透明度”计划, 目的在于增加人们对可重复性问题的认识, 增强研究者在研究设计要素方面的正规训练, 改善研究资助申请的评审, 增加对研究者的资助稳定度以使他们能够使用更合适和复杂的研究设计 (Tabak 2013)。该NIH指南计划的内容之一是开发出一套关于研究诚信的试验性训练模式, 因为研究诚信关乎实验偏倚和研究设计。活动目的在于为研究者提供特定的指南, 增加他们对可能影响研究发现可信性的研究实践的认识, 进而改善他们发表文章的质量。研究指南

也可用于改善研究资助申请和杂志的同行评审过程, 以保证对既往研究和即将开展的研究进行更为系统和严格的评价。

Hooijmans和Ritskes-Hoitinga (2013)最近发表了一项进展报告, 针对研究转化而开展的临床前期和动物实验性研究提出了一系列倡议, 试图解决研究可重复性问题。事实上, 动物实验性研究在除药物开发以外的许多领域都不可或缺。既往保护公众免受有害环境暴露的政策制定很大程度上依赖于动物实验性研究的结果。在更大范围的环境健康科学

研究领域, 重要的研究证据可以来源于各种不同设计的流行病学研究和机制研究。在公共卫生决策中使用一致并清晰的完整研究证据具有极为重大的全球重要性。

采取某些补救方法以改善对研究方法关键点的报告或许是解决问题的最简单方法, 这是因为有针对不同类型研究的报告质量检查清单可供使用。检查清单针对的研究类型包括临床试验 (Schulz et al. 2010)、观察性人群研究 (von Elm et al. 2008)、动物研究 (Hooijmans et al. 2010; Kilkenny et al. 2010) 和体外研究 (Schneider et al. 2009) (可参考EQUATOR Network 2014)。越来越多的杂志, 包括*Nature*系列杂志、*Plos One*和*Environmental Health Perspectives*, 目前已向作者提供明晰的指南以指导他们在稿件中对相应条目做出准确的报告。

系统综述的本质在于应用清晰的、严谨的、客观的和可重复的方法进行有依据的识别、选择、评价和综合相关研究的结果。系统综述方法在评估中的应用并不排斥对专家意见的需求或应用。这些方法有助于更为清楚地理解决策制定所依据的关键研究证据, 同时也能增加决策的可靠性。

建立可用于系统性评价研究设计和实施的合理性, 或者通俗地说研究质量的指南, 是极具挑战性的。虽然目前已有的一套合理的综合性方法用于评价人体临床试验的内部可靠性 (偏倚风险) (Higgins and Green 2011), 但是对于如何评价来源于人群观察性研究、动物实验性研究和体外研究的发现

和结论是否反映了研究的真实结果仍缺少类似的通行方法。系统综述领域的方法学发展或许能够为上述研究数据的处理提供合适的方法。

目前对系统综述方法的应用研究热点已从传统的人体临床试验领域拓展到毒理学和药理学等其他证据来源主体（涵盖于观察性研究、动物实验性研究和体外研究的范围内）（National Research Council 2009; Rooney et al. 2014; Sena et al. 2007; Woodruff and Sutton 2011）。例如，国家毒理学项目（National Toxicology Program, NTP）健康评价与转化办公室（Office of Health Assessment and Translation, OHAT）正在开展国际性合作，以开发出一种用于系统综述和证据整合的正规方法，该方法通过技术专家顾问咨询、科学顾问委员会、其他开展循证评价的机构或项目以及利益相关者的公开评述等方式对研究证据进行评估（Rooney et al. 2014）。导航指南工作组（Navigation Guide Work Group）已经建立了一个类似的框架，近期的案例研究也支持将系统综述方法应用于环境健康评估的可行性。进行系统综述的一个关键点是研究质量评估，包括研究的内部可靠性或偏倚风险。NTP、导航指南和其他机构开展的工作有助于开发出一种偏倚风险评价工具，适用于各种不同的人体、动物和机制性研究设计评价。该工具同时也有助于开发出新的评价和整理方法，以处理来源于不

同研究证据主体的数据。目前正在开发中的系统综述方法在某些方面有所不同，但实施方式上基本相似。OHAT开发出的灵活框架允许对合理开展的环境健康评价进行有针对性的评估。这些环境健康评价涵盖的信息可以来源于众多不同的研究类型和设计。上述框架可以根据实际使用情况进行不断的细化和完善。

今天对生物医学研究的投资必然对未来的类人生存质量改善起到推动作用。评价已发表科学发现的可重复性对保持生物医学研究的可信度至关重要。作者相信在整个科学研究过程中，包括在研究概念设计、研究资金申请和评审、研究开展、研究报告及最终将研究结果应用于环境健康科学或生物医学研究其他领域的结论导出等阶段，如果能广泛采用和遵守系统综述的各要素，将会显著改善公共卫生决策的准确性并提高科学投资的回报率。

*作者声称本文没有实际或潜在利益冲突。

译自EHP 122(7):A176-A177 (2014)

翻译：吴少伟

*本文参考文献请浏览英文原文

原文链接

<http://dx.doi.org/10.1289/ehp.1408671>

FRESH BACTERIA CASSEROLE. MMMM.

REFRIGERATE LEFTOVERS PROMPTLY, AND KEEP THE FRIDGE AT 40°F OR BELOW TO SLOW BACTERIA GROWTH.

CHILL



COOK



SEPARATE



CLEAN



KEEP YOUR FAMILY SAFER FROM FOOD POISONING
Check your steps at FoodSafety.gov

